Bernoulli **17**(2), 2011, 628–642
DOI: 10.3150/10-BEJ285

An approximate quantum Cramér–Rao bound based on skew information

ALESSANDRA LUATI

Department of Statistics, University of Bologna, via Belle Arti 41, 40126 Bologna, Italy.
E-mail: *alessandra.luati@unibo.it*

A closed-form expression for Wigner–Yanase skew information in mixed-state quantum systems is derived. It is shown that limit values of the mixing coefficients exist such that Wigner–Yanase information is equal to Helstrom information. The latter constitutes an upper bound for the classical expected Fisher information, hence the inverse Wigner–Yanase information provides an approximate lower bound to the variance of an unbiased estimator of the parameter of interest. The advantage of approximating Helstrom's sharp bound lies in the fact that Wigner–Yanase information is straightforward to compute, while it is often very difficult to obtain a feasible expression for Helstrom information. In fact, the latter requires the solution of an implicit second order matrix differential equation, while the former requires just scalar differentiation.

Keywords: Cramér–Rao-type bounds; Fisher information; parametric quantum models

1. Introduction

Classical statistics applied to quantum systems gives rise to more than one Fisher information quantity for the unknown parameter θ that specifies the state of the system, denoted by $\rho(\theta)$. The classical Fisher information $i(\theta, M)$ measures the precision of an unbiased estimator $t(x)$ of θ based on the outcome of an arbitrary measurement M, via the Cramér–Rao bound

$$\mathrm{Var}\{t(x)\} \geq i(\theta, M)^{-1}. \tag{1}$$

Various quantum analogues of classical information may be obtained directly based on quantum operators and without performing any measurement. The duality between classical and quantum information is a consequence of the interaction between the microscopic environment, where quantum systems evolve, and the macroscopic world, where measurements are performed. On the other hand, the existence of many quantum versions of one classical quantity is a characteristic feature of quantum mechanics, where states and measurements are represented by non-commutative operators or, in finite dimensions, by matrices.

This is an electronic reprint of the original article published by the ISI/BS in *Bernoulli*, 2011, Vol. 17, No. 2, 628–642. This reprint differs from the original in pagination and typographic detail.

1350-7265 © 2011 ISI/BS

Quantum analogs of classical information are derived according to which expression of Fisher information is generalized to the quantum setting, or to which operator-valued version of the logarithmic derivative is chosen, for example, the left or right logarithmic derivative ([27], [16], Section 6.6) or the symmetric logarithmic derivative ([14, 14], Section 8.4). As a result, each quantum information quantity inherits different properties and, consequently, may find application in different inferential problems, essentially related to Cramér–Rao-type bounds. Examples are the Helstrom [14] bound, where the information is based on the symmetric logarithmic derivative, and the two variants based on the left and right logarithmic derivatives considered by Belavkin [6]. Holevo [16] derived a bound based on vectors of matrix derivatives that can be attained asymptotically [11].

The relevance of the symmetric logarithmic derivative in the statistical literature is due to its relation with classical Fisher information. In fact, Helstrom information, $I_H(\theta)$, constitutes an upper bound for classical Fisher information [7],

$$i(\theta, M) \leq I_H(\theta), \tag{2}$$

from which the bound

$$\mathrm{Var}\{t(x)\} \geq I_H(\theta)^{-1} \tag{3}$$

[14] is obtained as a corollary. Since, in the quantum setting, the classical information formally depends on the measurement carried out in the system, much attention has been devoted in the literature to the search for measurements maximizing Fisher information. Major contributions have come from [4] and [18, 19], where necessary and sufficient conditions for equality between classical Fisher and quantum Helstrom information are derived, based on different hypotheses on the state of the quantum system.

Another strand of the literature has investigated the relations among quantum information quantities. Hayashi [13] compares the Kubo–Mori–Bogoliubov information ([1], Section 7.3) with the Helstrom information, applying to quantum estimation the large deviation viewpoint of Bahadur [2, 3]. Geometric relations based on the metric properties of quantum information are studied in [9, 12, 23, 24]. Recently, Luo [21] uncovered the relation between Helstrom and Wigner–Yanase [26] skew information, previously investigated by Luo [20] and also considered in [10], in the most basic quantum systems, called pure states. Specifically, Luo [21] showed that, in pure states, Wigner–Yanase information, $I_{WY}(\theta)$, is exactly twice the Helstrom information, $I_{WY}(\theta) = 2I_H(\theta)$, but left unsolved the problem of obtaining an exact solution in the generic mixed-state case which arises when mixtures of pure states are considered.

In this paper, we derive an explicit expression for Wigner–Yanase information in mixed-state models and relate it to both the Helstrom information obtained in the mixed state and the Wigner–Yanase information obtained in the pure states involved in the mixture. The connection between Wigner–Yanase and Helstrom information is explored in two-dimensional mixtures of orthogonal pure states, where a convenient expression for Helstrom information is available [18]. While in pure states, $I_H(\theta)$ and $I_{WY}(\theta)$ are equal up to a constant factor, in mixed states, their relation depends on the mixing coefficients. We show that for certain values of the mixing coefficients, $I_{WY}(\theta)$ is approximately equal

to $I_H(\theta)$. This suggests the possibility of approximating $I_H(\theta)$ by $I_{WY}(\theta)$ in the quantum Cramér–Rao bound (3), as well as in the information inequality (2). The advantage lies in the fact that Wigner–Yanase information can be easily calculated, unlike Helstrom information. In fact, while the latter arises as the solution of an implicit second order matrix differential equation, Wigner–Yanase information is simply obtained by scalar differentiation. Hence, the approximation provides a feasible bound.

The paper is organized as follows. Section 2 reviews the basics of quantum statistical inference and introduces the quantum analogs of classical expected Fisher information. Further details on probability and statistics in quantum systems can be found in the books by Holevo [16] and Helstrom [15] and in the paper by Barndorff-Nielsen, Gill and Jupp [5]; excellent references for quantum information theory are Nielsen and Chuang [22] and Petz [25]. The relation between Helstrom and Wigner–Yanase information in mixed states is derived in Section 3 and the generalizations are discussed in Section 4. Proofs and technical details are deferred to the Appendix. We conclude this introduction by observing that while quantum Cramér–Rao bounds are intended for vector parameter models, there is no general relation between the various quantum information inequalities when the parameter is a vector [5]. Hence, in the following, we will restrict our attention to one-parameter models.

2. Classical and quantum information

This paper is concerned with two quantum analogs of classical expected Fisher information that arise by generalizing to the quantum setting the following expressions

$$i(\theta, M) = \int_{\mathbb{G}_+} \left(\frac{\partial}{\partial\theta} \log p(x;\theta)\right)^2 p(x;\theta)\mu(\mathrm{d}x) \tag{4}$$

$$= 4\int_{\mathbb{G}_+} \left(\frac{\partial}{\partial\theta} \sqrt{p(x;\theta)}\right)^2 \mu(\mathrm{d}x), \tag{5}$$

where $X:(\Omega,\mathcal{F},P) \to (\mathbb{G},\mathcal{G},P_X)$ is a random variable characterized by the distribution $P_X(\cdot;\theta)$ with density $p(x;\theta)$ with respect to the σ-finite measure μ on $(\mathbb{G},\mathcal{G})$, $\mathbb{G}_+ = \{x \in \mathbb{G} : p(x;\theta) > 0\}$, $\theta \in \Theta \subset \mathbb{R}$ is the unknown parameter of interest and the usual regularity conditions apply [8]. The first equation bears a conceptual meaning, being the expected value of the square score function $l_{/\theta} = \frac{\partial}{\partial\theta} \log p(x;\theta)$, while the second is just an equivalent formulation.

If the random variable X describes an experiment in a quantum system, then its probability law depends on the state of the system, denoted by $\rho(\theta)$, as it is specified by the unknown parameter, and on the measurement M that is carried out in the system, in the following way (trace rule for probability):

$$P_X(G) = \mathrm{tr}\{\rho(\theta)M(G)\} \qquad \forall G \in \mathcal{G}.$$

More formally, $\rho(\theta)$ is a density matrix, that is, a self-adjoint, non-negative and trace-one linear operator acting on an a n-dimensional complex Hilbert space $\mathcal{H}_n$, and M is

a probability operator-valued measure, that is, a set of non-negative self-adjoint linear operators defined on the measure space $(\mathbb{G}, \mathcal{G})$ and taking values in $\mathcal{H}_n$, such that $M(\mathbb{G}) = \mathbf{I}$, the identity operator, $M(\varnothing) = \mathbf{O}$, the null operator, and $M(\bigcup_{h=1}^{\infty} G_h) = \sum_{h=1}^{\infty} M(G_h)$ if $G = \bigcup_{h=1}^{\infty} G_h, G_h \cap G_l = \varnothing$, for all $h, l = 1, \ldots, \infty$, $h \neq l$. Whenever $n < \infty$, the Hilbert space $\mathcal{H}_n$ can be identified with the n-dimensional Euclidean complex space $\mathbb{C}^n$ endowed with the standard inner product and it is equivalent to refer to self-adjoint operators or to Hermitian matrices. If M is absolutely continuous with respect to μ on $(\mathbb{G}, \mathcal{G})$, such that $M(G) = \int_G m(x)\mu(\mathrm{d}x)$ for all $G \in \mathcal{G}$, where $m(x)$ is non-negative and Hermitian, then $P_X(\cdot; \theta)$ is absolutely continuous with respect to μ and the density of X is

$$p(x;\theta) = \mathrm{tr}\{\rho(\theta) m(x)\}. \tag{6}$$

The expected value of the random variable X is

$$\mathrm{E}\{X\} = \int_{\mathbb{G}} x p(x;\theta)\mu(\mathrm{d}x) = \mathrm{tr}\left\{\rho(\theta) \int_{\mathbb{G}} x m(x)\mu(\mathrm{d}x)\right\}.$$

Combining (4) or (5) with (6), we obtain the classical expected Fisher information

$$i(\theta, M) = \int_{\mathbb{G}_+} (\mathrm{tr}\{\rho_{/\theta} m(x)\})^2 p(x;\theta)^{-1}\mu(\mathrm{d}x),$$

where $\rho_{/\theta}$ is the matrix whose ijth generic element is the derivative with respect to θ of the generic element of $\rho(\theta)$, that is, $[\rho_{/\theta}]_{ij} = \frac{\partial}{\partial\theta}[\rho(\theta)]_{ij}$.

Quantum analogs of classical information are obtained without performing any measurement on the system. The density matrix $\rho(\theta)$ plays the role of the density $p(x;\theta)$ and the expected value of any observable A, defined as a self-adjoint operator, is $\mathrm{E}\{A\} = \mathrm{tr}\{\rho(\theta)A\}$. Note that by defining $A := \int_{\mathbb{G}} x m(x)\mu(\mathrm{d}x)$, we have the formal connection with the expected value of the random variable X.

Helstrom [14] obtained the quantum information $I_H(\theta)$ by generalizing (4) to the operator-valued quantum setting via the symmetric logarithmic derivative, which is the self-adjoint operator $\rho_{/\!/\theta}$ implicitly defined by the relation

$$\rho_{/\theta} = \tfrac{1}{2}[\rho(\theta)\rho_{/\!/\theta} + \rho_{/\!/\theta}\rho(\theta)]. \tag{7}$$

The result is

$$I_H(\theta) = \mathrm{tr}\{\rho(\theta)\rho_{/\!/\theta}^2\}, \tag{8}$$

which, in fact, is the expected value of the (observable) square symmetric $\rho_{/\!/\theta}$. The quantity $\rho_{/\!/\theta}$ is also known as the *quantum score* since, besides (8), it satisfies

$$\mathrm{E}\{\rho_{/\!/\theta}\} = 0 \tag{9}$$

(see [5], page 789). The proof of the information inequality (2) for the one-dimensional parameter case is based on the Cauchy–Schwarz inequality with Hilbert–Schmidt inner

product. The bound (3) can be obtained as a consequence of (2) or, directly, by following the derivation of the classic Cramér–Rao bound.

On the other hand, Wigner–Yanase [26] information, denoted by $I_{WY}(\theta)$, is obtained by generalizing (5) in a straightforward manner:

$$I_{WY}(\theta) = 4\,\mathrm{tr}\{[(\rho(\theta)^{1/2})_{/\theta}]^2\}. \tag{10}$$

Luo [21] derived a relation between $I_{WY}(\theta)$ and $I_H(\theta)$ in the case where the system is described by a pure state density matrix of the form $\rho(\theta) = |\psi(\theta)\rangle\langle\psi(\theta)|$, where $|\psi(\theta)\rangle$ is a unit vector in $\mathcal{H}_n$ and we have used the Dirac notation according to which $|\psi(\theta)\rangle$ denotes a column vector and $\langle\psi(\theta)|$ is its Hermitian transpose. Specifically, the relation is

$$I_{WYp}(\theta) = 2I_{Hp}(\theta), \tag{11}$$

where the attached subscript p denotes here that the quantities are associated with a pure state. The proof can be shortened by observing that, in pure states, $\rho(\theta)^{1/2} = \rho(\theta)$ and

$$I_{Hp}(\theta) = 2\,\mathrm{tr}\{(\rho_{/\theta})^2\}$$

(see [19], equation (3.1)), from which (11) clearly follows by direct comparison with (10).

According to quantum state estimation theory, pure states represent the best knowledge one can have about some specific properties of the system under observation. Mixed states, obtained as convex combinations of pure states, indicate a situation of partial knowledge of the system. They represent probabilistic mixtures, in the sense that the system under observation is in the state $\rho_i(\theta)$ with probability $w_i(\theta)$, $i = 1, \ldots, m$, and $\sum_{i=1}^m w_i(\theta) = 1$.

To obtain a relation between Helstrom and Wigner–Yanase information in mixed states, it is necessary to express both of the information quantities as functions of common variables (up to some transformation), such as, for example, the corresponding information quantities in the mixing pure states. In this setting, the main difficulty is related to Helstrom information since explicit solutions of (7) when $\rho(\theta)$ is a mixed state are not usually available in a convenient form. Restricting to mixtures of two-dimensional orthogonal pure states allows this difficulty to be overcome and also a suitable expression for skew information to be derived.

Two-dimensional systems play a crucial role in quantum mechanics. Electrons, qubits and spin-$\frac{1}{2}$ particles are just some examples of systems in $\mathbb{C}^2$. In addition, two-dimensional mixtures of orthogonal states have an appealing geometric interpretation, based on the Bloch (or Poincaré, or Riemann) sphere representation of states in two-dimensional complex Hilbert spaces by means of unit vectors in the real three-dimensional Euclidean space. In fact, if $\mathcal{H}_n = \mathbb{C}^2$, then the set of pure states is the surface of the unit sphere and the set of mixed states is the interior of the corresponding unit ball (see [25]). Mixtures of two pure states can be represented as points in the interior of the sphere, on the straight line joining the two points on the surface. If the generating pure states are orthogonal (opposite on the sphere), then the corresponding mixed states lie on the diameters of the great circles and therefore the set of such states with given weights can

be represented by the spheres embedded in the unit sphere with the same center, but with radius less than one and dependent on the weights of the mixtures. Specifically, we shall consider all of the spheres embedded in the unit sphere with the same center and radius equal to $|2w(\theta)-1|$, where $w(\theta)\in(0,1)$ determines the weights of the mixture.

3. Skew information in mixed states

A one-parameter two-dimensional mixed state can be represented as

$$\rho(\theta)=w(\theta)\rho_1(\theta)+(1-w(\theta))\rho_2(\theta), \tag{12}$$

where $\rho_1(\theta)=|\psi_1(\theta)\rangle\langle\psi_1(\theta)|$ and $\rho_2(\theta)=|\psi_2(\theta)\rangle\langle\psi_2(\theta)|$ are orthogonal pure states, in the sense that $\langle\psi_1(\theta)|\psi_2(\theta)\rangle=0$ and $w(\theta)$ is a function of θ taking values in the real interval $(0,1)$, $w(\theta)\neq\frac{1}{2}$. Note that orthogonal pure states are such that $\rho_h(\theta)\rho_h(\theta)=\rho_h(\theta)$ and $\rho_h(\theta)\rho_k(\theta)=\mathbf{O}$ for $h\neq k$. Let us denote by $\rho_{/\theta h}$ and $\rho_{/\!/\theta h}$ the term-by-term first derivative and the symmetric logarithmic derivative of $\rho_h(\theta), h=1,2$, with respect to θ, respectively, and by $I_{Hh}(\theta)$ and $I_{WYh}(\theta)$, the Helstrom and Wigner–Yanase information extracted from the pure state $\rho_h(\theta)$, respectively. We assume that $w(\theta)$ has continuous first derivative, $w_{/\theta}$, such that $w_{/\theta}\to 0$ faster than $\sqrt{w(\theta)}$ and $\sqrt{1-w(\theta)}$ for all θ such that $w(\theta)\to 0$ and $w(\theta)\to 1$, respectively. With these premises, we can state the following proposition.

Proposition 1. *In the mixed state (12), Wigner–Yanase skew information is*

$$I_{WY}(\theta)=\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}+\left(1-2\sqrt{w(\theta)(1-w(\theta))}\right)I_{WY1}(\theta) \tag{13}$$

and the following relation holds

$$I_{WY}(\theta)=\alpha_w(\theta)I_H(\theta)+\beta_w(\theta), \tag{14}$$

where

$$\alpha_w(\theta)=\frac{2}{1+2\sqrt{w(\theta)(1-w(\theta))}} \quad \textit{and} \quad \beta_w(\theta)=-\frac{1-2\sqrt{w(\theta)(1-w(\theta))}}{1+2\sqrt{w(\theta)(1-w(\theta))}}\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}.$$

The proof is in Section A.1 and, concerning (13), it is based on some properties of pure states and quantum information quantities. The proof of (14) is based on the comparison with Helstrom information derived in [18], Lemma 2, where a specific choice of $|\psi_1(\theta)\rangle$ and $|\psi_2(\theta)\rangle$ was made in order to derive the symmetric logarithmic derivative of the mixed state as a function of the symmetric logarithmic derivative of $\rho_1(\theta)$,

$$I_H(\theta)=\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}+(2w(\theta)-1)^2 I_{H1}(\theta). \tag{15}$$

The boundary conditions on $w_{/\theta}$ ensure that $I_{WY}(\theta) \to I_{WY1}(\theta)$ and $I_H(\theta) \to I_{H1}(\theta)$ when $w(\theta) \to 0, 1$ and imply that $\beta_w(\theta) \to 0$ when $w(\theta) \to 0, 1$. Hence, they encompass the case where the mixing coefficient does not depend on θ ($w_{/\theta} = 0 \ \forall\theta \in \Theta \subset \mathbb{R}$).

In the analysis of the relation between $I_{WY}(\theta)$ and $I_H(\theta)$, we first consider the latter case, that is, w does not depend on θ, which implies that

$$I_{WY}(\theta) = \alpha_w I_H(\theta),$$

where α_w is function of w, symmetric with respect to $w = \frac{1}{2}$, where it reaches its minimum of 1, and with maximum of 2 at $w = 0$ and $w = 1$. The implications are evident: when the mixing coefficients tend to 0 and 1, the skew information tends to be twice the Helstrom information since the mixture tends to reproduce a pure state system. Conversely, when $w \to \frac{1}{2}$, the distance between the two quantum information quantities reaches its minimum, in that they tend to be equal. The limit case $w \to \frac{1}{2}$ is of special interest to us, so we shall return to it in the discussion of Proposition 1 for the case where w depends on θ. Furthermore, both $I_{WY}(\theta)$ and $I_H(\theta)$ are smaller than the corresponding information quantities in the pure states. This is coherent with quantum theory, according to which pure states represent the best knowledge that one can have about a quantum system.

On the other hand, if the mixing coefficients depend on θ, then a sufficient condition for $I_{WY}(\theta) < I_{WY1}(\theta)$ and $I_H(\theta) < I_{H1}(\theta)$ is that

$$I_{WY1}(\theta) > \frac{(w_{/\theta})^2}{2w^2(\theta)(1 - w(\theta))^2},$$

which follows by simple algebraic manipulation of (13) and (15), together with the identity $I_{WY1}(\theta) = 2I_{H1}(\theta)$. In other words, it is sufficient that the condition $I_H(\theta) < I_{H1}(\theta)$ is satisfied to ensure that $I_{WY}(\theta) < I_{WY1}(\theta)$.

The limit cases $w(\theta) \to 0, \frac{1}{2}, 1$ turn out to be identical to the limit cases when w is constant. In particular, if $w(\theta) \to \frac{1}{2}$, then $\beta_w(\theta) \to 0$ and $\alpha_w(\theta) \to 1$ so that we find $I_{WY}(\theta) \to I_H(\theta)$. Hence, we can state the following corollary to Proposition 1. This is a direct consequence of the quantum Cramér–Rao bound (3) and hence the proof is omitted.

Corollary 1. *In the mixed state (12), if* $w(\theta) \to \frac{1}{2}$*, then* $I_{WY}(\theta) \to I_H(\theta)$ *and* $I_{WY}(\theta)^{-1}$ *constitutes an approximate lower bound for* $\mathrm{Var}\{t(x)\}$.

The equality between $I_H(\theta)$ and $I_{WY}(\theta)$ is a limit condition that holds in a degenerate case. Precisely, if $w(\theta) = \frac{1}{2}$, then $\rho(\theta) = \frac{1}{2}\mathbf{I}$. This case (the center of the unit sphere) represents the maximum entropy situation, that is, complete ignorance about the quantum system under observation. Nevertheless, the bound has the following interpretation: the precision of the feasible Wigner–Yanase quantum Cramér–Rao bound increases as long as the mixture approaches the maximum entropy case. Values of $w(\theta)$ that lie in a neighborhood of $w(\theta) = \frac{1}{2}$ give rise to sensible mixtures and feasible approximated quantum Cramér–Rao bounds. In the two-dimensional case, the size of the approximation can be measured using equation (14).

In pure states, where we are equipped with all the quantum information given by the system, knowledge of I_{WY} is equivalent to knowledge of I_H (up to a constant factor, they are equal); in mixed states, the distance between the two quantum information quantities depends on the probability of being in a (one-dimensional) space rather than on its orthogonal complement. The same distance between $I_H(\theta)$ and $I_{WY}(\theta)$ decreases as long as the knowledge of the system diminishes, eventually tending to zero in the maximum entropy case (uniform distribution on the coefficients).

4. Discussion

We derived the relation between Helstrom and Wigner–Yanase information in two-dimensional mixed-state systems, where a geometric interpretation of the states as spheres embedded in the unit sphere can be drawn. We now discuss the case of general mixed states

$$\rho(\theta) = \sum_{i=1}^{m} w_i(\theta)\rho_i(\theta),$$

where the $\rho_i(\theta)$ are pure states defined on $\mathbb{C}^n$ and not necessarily orthogonal, and

$$\sum_{i=1}^{m} w_i(\theta) = 1.$$

If the mixed state admits the spectral decomposition

$$\rho(\theta) = \sum_{l=1}^{n} \lambda_l(\theta)\rho_l(\theta), \tag{16}$$

then $I_{WY}(\theta)$ can be obtained directly, following the proof of Proposition 1, as

$$I_{WY}(\theta) = \sum_{l=1}^{n} \lambda_l(\theta) I_{WY,l}(\theta) + \sum_{l=1}^{n} \frac{(\lambda_{/\theta l})^2}{\lambda_l(\theta)} + 4\sum_{l=1}^{n}\sum_{k\neq l} \sqrt{\lambda_l(\theta)\lambda_k(\theta)}\,\mathrm{tr}\{\rho_{/\theta l}\rho_{/\theta k}\}. \tag{17}$$

When $n = 2$, $\lambda_1(\theta) = w(\theta)$, $\lambda_2(\theta) = 1 - w(\theta)$, $(\lambda_{/\theta 1})^2 = (\lambda_{/\theta 2})^2 = (w_{/\theta})^2$ and (17) is equal to (13). On the other hand, concerning $I_H(\theta)$, a solution of (7) based on the spectral decomposition of $\rho(\theta)$ is [24]

$$\rho_{/\!/\theta} = \sum_{l=1}^{n}\sum_{k=1}^{n} \frac{2}{\lambda_l(\theta) + \lambda_k(\theta)} \rho_j(\theta)\rho_{/\theta}\rho_k(\theta), \tag{18}$$

which allows us to obtain the maximum Fisher information attainable in the mixed state (16),

$$I_H(\theta) = \sum_{l=1}^{n} \frac{(\lambda_{/\theta l})^2}{\lambda_l(\theta)} + \sum_{l=1}^{n}\sum_{k\neq l}\sum_{z=1}^{n} \frac{4\lambda_l(\theta)(\lambda_k(\theta) - \lambda_l(\theta))\lambda_z(\theta)}{(\lambda_l(\theta) + \lambda_k(\theta))^2}\,\mathrm{tr}\{\rho_l(\theta)\rho_{/\theta k}\rho_{/\theta z}\}. \tag{19}$$

The derivation is not direct and is therefore deferred to Section A.2 of the Appendix. It is easy to verify that when $n = 2$, (19) is equal to (15), since, besides $\lambda_1(\theta) = w(\theta)$, $\lambda_2(\theta) = 1 - w(\theta)$ and $(\lambda_{/\theta 1})^2 = (\lambda_{/\theta 2})^2 = (w_{/\theta})^2$, $4\,\mathrm{tr}\{\rho_l(\theta)\rho_{/\theta k}\rho_{/\theta z}\}$ is equal to plus or minus $I_{H1}(\theta)$, according to whether $z = k$ or $z \neq k$, respectively, as follows by the properties of two orthogonal pure states in two-dimensional spaces, $\rho_{/\!/\theta h} = 2\rho_{/\theta h}$, $h = 1, 2$, $\rho_{/\theta h} = -\rho_{/\theta k}$, $h \neq k$ and $I_{H2}(\theta) = I_{H1}(\theta)$ ([18], equations (i), (vi) and (ix), respectively).

It is evident that the two information quantities (17) and (19) are sensibly comparable only under some specific assumptions on the state of the quantum system, like those we have made in Section 3. Moreover, the two expressions do not explicitly depend on the mixing coefficients, nor on the mixing pure states, so their formal equivalence loses its interpretation. Nevertheless, rewriting $I_{WY}(\theta)$ in (17) as a function of $I_H(\theta)$ in (19) provides an eigenvalue-based condition of equality between the two quantum information quantities, one which allows us to derive a result equivalent to Corollary 1 in Section 3, that is, an approximate quantum Cramér–Rao bound based on skew information in n-dimensional mixed states that admit the spectral decomposition (16). We state this finding in a proposition, proved in Section A.3 of the Appendix, and a corollary, whose proof is omitted since it follows from the proposition in a straightforward manner.

Proposition 2. *In the mixed state (16),*

$$I_{WY}(\theta) = I_H(\theta) + \gamma_{\lambda,\rho}(\theta),$$

where, omitting the arguments of the functions,

$$\gamma_{\lambda,\rho} = -4\sum_{l=1}^{n}\sum_{k\neq l}\left((\lambda_l - \sqrt{\lambda_l\lambda_k})\,\mathrm{tr}\{\rho_{/l}\rho_{/k}\} + \sum_{z=1}^{n}\frac{\lambda_l(\lambda_k-\lambda_l)\lambda_z}{(\lambda_l+\lambda_k)^2}\,\mathrm{tr}\{\rho_l\rho_{/k}\rho_{/z}\}\right). \tag{20}$$

The proof essentially consists of rewriting Wigner–Yanase information in a convenient way and then comparing it with Helstrom information. The quantity $\gamma_{\lambda,\rho}(\theta)$ depends both on the eigenvalues of $\rho(\theta)$, playing the role of the mixing coefficients in the representation (16), and on the states $\rho_l(\theta)$, via their derivatives. Hence, it would be misleading to interpret $\gamma_{\lambda,\rho}(\theta)$ as a pure additive quantity, be it positive or negative. As a matter of fact, the dependence of $\gamma_{\lambda,\rho}(\theta)$ on $\lambda_l(\theta)$ and $\rho_l(\theta)$ implies that the quantity $\gamma_{\lambda,\rho}(\theta)$ can involve, at least, combinations of information quantities in the one-dimensional pure states that span the mixed-state system (16). To illustrate this, we shall consider the two-dimensional case, where $\gamma_{\lambda,\rho}(\theta) = (1-2\sqrt{w(\theta)(1-w(\theta))})^2 I_{H1}(\theta) = (\alpha_w(\theta)-1)I_H(\theta) + \beta_w(\theta)$. Unlike $\gamma_{\lambda,\rho}(\theta)$, the quantities $\alpha_w(\theta)$ and $\beta_w(\theta)$ depend only on the mixing coefficients and their interpretation is unequivocal. Also, note that in the two-dimensional case, $\gamma_{\lambda,\rho}(\theta) \to 0$ for all θ such that $w(\theta) \to \frac{1}{2}$, this leading to the approximate quantum Cramér–Rao bound stated in Corollary 1.

Corollary 1 can be generalized to the n-dimensional case by means of Proposition 2. In fact, it is immediate to see that in (20), if $\lambda_k(\theta) \to \lambda_l(\theta)$ for all k and l, then $\gamma_{\lambda,\rho}(\theta) \to 0$ and $I_{WY(\theta)} \to I_H(\theta)$. A sufficient condition for all of the λ to be equal is that $\lambda_l(\theta) \to \frac{1}{n}$ for all $l = 1, \ldots, n$ and the following corollary can be established as a consequence of the quantum Cramér–Rao bound (3).

Corollary 2. *In the mixed state (16), if* $\lambda_l(\theta) \to \frac{1}{n}$ *for all* $l = 1, 2, \ldots, n$*, then* $I_{WY}(\theta) \to I_H(\theta)$ *and* $I_{WY}(\theta)^{-1}$ *constitutes an approximate lower bound for* $\mathrm{Var}\{t(x)\}$.

We can therefore conclude that in a mixed state specified by its spectral decomposition, a limit condition analog to the one holding in two-dimensional systems holds as well. The interpretation of the approximate quantum Cramér–Rao bound specified in Corollary 2 is the same as in the two-dimensional case. As long as, in (16), the mixing coefficients tend to be uniformly distributed over n, Wigner–Yanase skew information tends to be equal to Helstrom information and can serve as an approximation of the latter in the information inequality (2) and in the quantum Cramér–Rao bound (3). The uniform distribution of the coefficients is a limit condition that holds in the degenerate maximum entropy case, but situations where the limit is approached give rise to sensible mixtures and feasible approximated quantum Cramér–Rao bounds. Deriving equations (17) and (19) shows that it is, in general, much easier to obtain a feasible expression for Wigner–Yanase information than for Helstrom information, which further justifies the use of (10) as an approximate upper bound for the classical expected Fisher information.

Appendix

A.1. Proof of Proposition 1

We derive $I_{WY}(\theta) = 4\,\mathrm{tr}\{[(\rho(\theta)^{1/2})_{/\theta})]^2\}$ in the mixed state (12). Since $|\psi_1(\theta)\rangle$ and $|\psi_2(\theta)\rangle$ are orthogonal and $\rho_1(\theta)$ and $\rho_2(\theta)$ are pure states, we have

$$\rho(\theta)^{1/2} = \sqrt{w(\theta)}\rho_1(\theta) + \sqrt{1 - w(\theta)}\rho_2(\theta). \tag{21}$$

Note that $\rho(\theta)^{1/2} \to \rho_h(\theta)$, $h = 2, 1$, if $w(\theta) \to 0, 1$. Also, note that $\rho(\theta)^{1/2}$ is not a density matrix since its trace is not equal to 1. The choice of $\rho(\theta)^{1/2}$, as in (21), as opposed its alternative with the negative sign, guarantees that the square root of the positive semidefinite self-adjoint operator $\rho(\theta)$ is positive semidefinite for all of the values $w(\theta)$ ([17], Theorem 2.6, page 405).

The elementwise derivative of (21) with respect to θ is given by

$$(\rho(\theta)^{1/2})_{/\theta} = \frac{w_{/\theta}}{2\sqrt{w(\theta)}}\rho_1(\theta) + \sqrt{w(\theta)}\rho_{/\theta 1} - \frac{w_{/\theta}}{2\sqrt{1 - w(\theta)}}\rho_2(\theta) + \sqrt{1 - w(\theta)}\rho_{/\theta 2},$$

where we have assumed that $w_{/\theta}$ tends to zero faster than $\sqrt{w(\theta)}$ and $\sqrt{1 - w(\theta)}$ for all θ such that $w(\theta) \to 0, 1$, respectively.

In taking the square and then the trace of the above quantity, we observe that the spectral theorem in $\mathbb{C}^2$ implies that $\rho_2(\theta) = \mathbf{I} - \rho_1(\theta)$ and, consequently, that $\rho_{/\theta 2} = -\rho_{/\theta 1}$, $\rho_{/\theta 2}\rho_{/\theta 1} = -(\rho_{/\theta 1})^2$ and $(\rho_{/\theta 2})^2 = (\rho_{/\theta 1})^2 = ((\rho_1(\theta)^{1/2})_{/\theta})^2$. Furthermore,

$$\mathrm{tr}\{\rho_k(\theta)\rho_{/\theta h}\} = 0 \qquad \forall h, k. \tag{22}$$

In fact, when $h=k$, equation (22) follows from the fact that, in pure states, $\rho_{/\!/\theta}=2\rho_{/\theta}$ ([19], equation (3.1)), combined with the fact that $\mathrm{E}\{\rho_{/\!/\theta}\}=2\mathrm{E}\{\rho_{/\theta}\}=0$, stated in (9). On the other hand, when $h\neq k$, using the definition of symmetric logarithmic derivative (7), $\mathrm{tr}\{\rho_k(\theta)\rho_{/\theta h}\}=\frac{1}{2}\,\mathrm{tr}\{(\rho_k(\theta)\rho_h(\theta)\rho_{/\!/\theta h}+\rho_k(\theta)\rho_{/\!/\theta h}\rho_h(\theta))\}=0$.

Hence,

$$\begin{aligned}
4\,\mathrm{tr}\{[(\rho(\theta)^{1/2})_{/\theta}]^2\} &= 4\left[\left(\frac{w_{/\theta}}{2\sqrt{w(\theta)}}\right)^2+\left(\frac{w_{/\theta}}{2\sqrt{1-w(\theta)}}\right)^2\right] \\
&\quad + 4(w(\theta)+1-w(\theta)-2\sqrt{w(\theta)}\sqrt{1-w(\theta)})\,\mathrm{tr}\{(\rho_{/\theta 1})^2\} \\
&= \frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}+(1-2\sqrt{w(\theta)(1-w(\theta))})I_{WY1}(\theta).
\end{aligned}$$

This completes the first part of the proof, that is, equation (13).

We now prove equation (14). Without loss of generality, we can choose

$$|\psi_2(\theta)\rangle = 2\sqrt{2}I_{WY1}^{-1/2}(\theta)\rho_{/\theta 1}|\psi_1(\theta)\rangle = I_{H1}^{-1/2}(\theta)\rho_{/\theta 1}|\psi_1(\theta)\rangle.$$

This setting is convenient because Lemma 2 in [18] applies, which states that in two-dimensional orthogonal mixed states where $|\psi_1(\theta)\rangle$ and $|\psi_2(\theta)\rangle$ are defined as above, Helstrom information is given by equation (15), which can be written as

$$(2w(\theta)-1)^{-2}\left[I_H(\theta)-\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}\right]=I_{H1}(\theta).$$

It follows from (11) and (13) that

$$I_{WY}(\theta)=\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}+2(1-2\sqrt{w(\theta)(1-w(\theta))})I_{H1}(\theta).$$

Combining the two former equations, we get

$$\begin{aligned}
I_{WY}(\theta) &= \frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}\left[1-\frac{2(1-2\sqrt{w(\theta)(1-w(\theta))}}{(2w(\theta)-1)^2}\right] \\
&\quad + \frac{2(1-2\sqrt{w(\theta)(1-w(\theta))}}{(2w(\theta)-1)^2}I_H(\theta) \\
&= \frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}\left[\frac{-(1-2\sqrt{w(\theta)(1-w(\theta))})^2}{(2w(\theta)-1)^2}\right] \\
&\quad + \frac{2(1-2\sqrt{w(\theta)(1-w(\theta))})}{(2w(\theta)-1)^2}I_H(\theta).
\end{aligned}$$

Given that $(2w(\theta)-1)^2 = (1-2\sqrt{w(\theta)(1-w(\theta))})(1+2\sqrt{w(\theta)(1-w(\theta))})$, we obtain

$$I_{WY}(\theta) = -\frac{(w_{/\theta})^2}{w(\theta)(1-w(\theta))}\frac{1-2\sqrt{w(\theta)(1-w(\theta))}}{1+2\sqrt{w(\theta)(1-w(\theta))}} + \frac{2}{1+2\sqrt{w(\theta)(1-w(\theta))}} I_H(\theta).$$

A.2. Proof of equation (19)

It follows from (16) and (18) that

$$\begin{aligned} I_H(\theta) &= \mathrm{tr}\{\rho(\theta)\rho^2_{/\!/\theta}\} \\ &= \mathrm{tr}\left\{\sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n}\frac{4\lambda_l(\theta)}{(\lambda_l(\theta)+\lambda_k(\theta))(\lambda_k(\theta)+\lambda_j(\theta))}\rho_l(\theta)\rho_{/\theta}\rho_k(\theta)\rho_{/\theta}\rho_j(\theta)\right\}. \end{aligned}$$

The linearity and the cyclical property of the trace operator, associated with the orthogonality of pure states $\rho_l(\theta)$ (by the spectral theorem, all of these operators are orthogonal projections onto one-dimensional supspaces of $\mathbb{C}^n$), give

$$I_H(\theta) = \sum_{l=1}^{n}\sum_{k=1}^{n}\frac{4\lambda_l(\theta)}{(\lambda_l(\theta)+\lambda_k(\theta))^2}\,\mathrm{tr}\{\rho_l(\theta)\rho_{/\theta}\rho_k(\theta)\rho_{/\theta}\}$$

and, omitting all of the arguments,

$$\begin{aligned} I_H(\theta) &= \sum_{l=1}^{n}\sum_{k=1}^{n}\frac{4\lambda_l}{(\lambda_l+\lambda_k)^2}\,\mathrm{tr}\left\{\rho_l\sum_{m=1}^{n}(\rho_m\lambda_{/m}+\rho_{/m}\lambda_m)\rho_k\sum_{z=1}^{n}(\rho_z\lambda_{/z}+\rho_{/z}\lambda_z)\right\} \\ &= \sum_{l=1}^{n}\sum_{k=1}^{n}\frac{4\lambda_l}{(\lambda_l+\lambda_k)^2}\,\mathrm{tr}\left\{\left(\rho_l\lambda_{/l}+\sum_{m=1}^{n}\lambda_m\rho_l\rho_{/m}\right)\left(\rho_k\lambda_{/k}+\sum_{z=1}^{n}\lambda_z\rho_k\rho_{/z}\right)\right\} \\ &= \sum_{l=1}^{n}\frac{(\lambda_{/l})^2}{\lambda_l} + \sum_{l=1}^{n}\frac{\lambda_{/l}}{\lambda_l}\sum_{z=1}^{n}\lambda_z\,\mathrm{tr}\{\rho_l\rho_{/z}\} + \sum_{l=1}^{n}\frac{\lambda_{/l}}{\lambda_l}\sum_{m=1}^{n}\lambda_m\,\mathrm{tr}\{\rho_l\rho_{/m}\} \\ &\quad + \sum_{l=1}^{n}\sum_{k=1}^{n}\frac{4\lambda_l}{(\lambda_l+\lambda_k)^2}\,\mathrm{tr}\left\{\sum_{m=1}^{n}\sum_{z=1}^{n}\lambda_m\lambda_z\rho_l\rho_{/m}\rho_k\rho_{/z}\right\} \\ &= \sum_{l=1}^{n}\frac{(\lambda_{/l})^2}{\lambda_l} + 2\sum_{l=1}^{n}\frac{\lambda_{/l}}{\lambda_l}\sum_{z=1}^{n}\lambda_z\,\mathrm{tr}\{\rho_l\rho_{/z}\} \\ &\quad + \sum_{l=1}^{n}\sum_{k=1}^{n}\frac{4\lambda_l}{(\lambda_l+\lambda_k)^2}\,\mathrm{tr}\left\{\sum_{m=1}^{n}\sum_{z=1}^{n}\lambda_m\lambda_z\rho_l\rho_{/m}\rho_k\rho_{/z}\right\}. \end{aligned}$$

The second term of the above equation is null; see (22). In order to simplify the final expression of $I_H(\theta)$, it is convenient to analyze the implications of the product $\rho_l \rho_{/m}$ for the third term of the above equation. When $l = m$, we can prove that

$$\operatorname{tr}\{\rho_l \rho_{/l} \rho_l \rho_{/l}\} = 0 \qquad \text{for all } l = 1, \ldots, n. \tag{23}$$

By definition of pure state, $\rho_l \rho_l = \rho_l$ and, consequently, $\rho_{/l}\rho_l + \rho_l \rho_{/l} = \rho_{/l}$, that is, $\rho_l \rho_{/l} = \rho_{/l} - \rho_{/l}\rho_l$, which, substituted into $\operatorname{tr}\{\rho_l \rho_{/l} \rho_l \rho_{/l}\}$, gives $\operatorname{tr}\{(\rho_{/l} - \rho_{/l}\rho_l)\rho_l \rho_{/l}\} = 0$. On the other hand, let us consider any $m \neq l$. It is straightforward to prove that

$$\rho_l \rho_{/m} = -\rho_{/l}\rho_m, \qquad m \neq l. \tag{24}$$

In fact, from the spectral theorem, $\rho_l \rho_m = \mathbf{O} \Rightarrow \rho_{/l}\rho_m + \rho_l \rho_{/m} = \mathbf{O} \Rightarrow \rho_l \rho_{/m} = -\rho_{/l}\rho_m$, that is, non-null terms arise for $m \neq l$ when $m = k = z$.

Let us now consider the third term of the last expression for $I_H(\theta)$. First, note that if $k = l$, then the trace is null. In fact, if $m = l$, then $\operatorname{tr}\{\rho_l \rho_{/m} \rho_l \rho_{/z}\} = \operatorname{tr}\{\rho_l \rho_{/l} \rho_l \rho_{/z}\} = 0$ both if $z = l$, by (23), and if $z \neq l$, by (24), and by the cyclical properties of the trace operator; on the other hand, if $m \neq l$, then $\operatorname{tr}\{\rho_l \rho_{/m} \rho_l \rho_{/z}\} = \operatorname{tr}\{-\rho_{/l}\rho_m \rho_l \rho_{/z}\} = 0$ for all z.

Therefore, it follows that $I_H(\theta)$ can be simplified as

$$\begin{aligned}
I_H(\theta) &= \sum_{l=1}^{n} \frac{(\lambda_{/l})^2}{\lambda_l} + \sum_{l=1}^{n} \sum_{k \neq l} \frac{4\lambda_l}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\left\{ \sum_{m=1}^{n} \sum_{z=1}^{n} \lambda_m \lambda_z \rho_l \rho_{/m} \rho_k \rho_{/z} \right\} \\
&= \sum_{l=1}^{n} \frac{(\lambda_{/l})^2}{\lambda_l} + \sum_{l=1}^{n} \sum_{k \neq l} \frac{4\lambda_l}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\left\{ \sum_{m \neq k} \sum_{z=1}^{n} \lambda_m \lambda_z \rho_l \rho_{/m} \rho_k \rho_{/z} \right\} \\
&\quad + \sum_{l=1}^{n} \sum_{k \neq l} \frac{4\lambda_l}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\left\{ \sum_{z=1}^{n} \lambda_k \lambda_z \rho_l \rho_{/k} \rho_k \rho_{/z} \right\},
\end{aligned}$$

where we have distinguished the case $m \neq k$ from the case $m = k$. Now, note that, when $m \neq k$, the trace in the second summand is null unless $m = l$, as follows by (24). Using the same arguments, but applied to $l \neq k$ in the third summand, we can simplify the final expression of Helstrom information in a generic mixed state as

$$\begin{aligned}
I_H(\theta) &= \sum_{l=1}^{n} \frac{(\lambda_{/l})^2}{\lambda_l} + \sum_{l=1}^{n} \sum_{k \neq l} \frac{4(\lambda_l)^2}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\left\{ \sum_{z=1}^{n} \lambda_z (-\rho_l \rho_{/k} \rho_{/z}) \right\} \\
&\quad + \sum_{l=1}^{n} \sum_{k \neq l} \frac{4\lambda_l \lambda_k}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\left\{ \sum_{z=1}^{n} \lambda_z (-\rho_{/l} \rho_k \rho_{/z}) \right\} \\
&= \sum_{l=1}^{n} \frac{(\lambda_{/l})^2}{\lambda_l} + \sum_{l=1}^{n} \sum_{k \neq l} \sum_{z=1}^{n} \frac{4\lambda_l (\lambda_k - \lambda_l) \lambda_z}{(\lambda_l + \lambda_k)^2} \operatorname{tr}\{\rho_l \rho_{/k} \rho_{/z}\}.
\end{aligned}$$

A.3. Proof of Proposition 2

Let us consider equation (17). We observe that $I_{WY,l}(\theta) = 4\,\mathrm{tr}\{\rho_{/\theta l}\rho_{/\theta l}\} = -4\sum_{k\neq l}\mathrm{tr}\{\rho_{/\theta l}\rho_{/\theta k}\}$, as follows by $\rho_{/\theta l} = -\sum_{k\neq l}\rho_{/\theta k}$, a consequence of the spectral identity $\sum_{l=1}^{n}\rho_l(\theta) = \mathbf{I}$. Hence,

$$I_{WY}(\theta) = \sum_{l=1}^{n}\frac{(\lambda_{/\theta l})^2}{\lambda_l(\theta)} - 4\sum_{l=1}^{n}\sum_{k\neq l}\left(\lambda_l(\theta) - \sqrt{\lambda_l(\theta)\lambda_k(\theta)}\right)\mathrm{tr}\{\rho_{/\theta k}\rho_{/\theta l}\}.$$

By replacing, in the above equation, the first term on the right-hand side by the same quantity obtained from (19) as a function of $I_H(\theta)$, and rearranging, we obtain the claimed result.

Acknowledgements

I would like to thank the Editor and a referee for their comments.

References

[1] Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry*. Oxford: Oxford Univ. Press. MR1800071

[2] Bahadur, R.R. (1960). On the asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252. MR0293767

[3] Bahadur, R.R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324. MR0207085

[4] Barndorff-Nielsen, O.E. and Gill, R.D. (2000). Fisher information in quantum statistics. *J. Phys. A* **33** 4481–4490. MR1768745

[5] Barndorff-Nielsen, O.E., Gill, R.D. and Jupp, P.E. (2003). On quantum statistical inference. *J. Roy. Statist. Assoc. B* **65** 775–816. MR2017871

[6] Belavkin, V.P. (1976). Generalized uncertainty relations and efficient measurements in quantum systems. *Teoret. Mat. Fiz.* **26** 316–329. MR0449358

[7] Braunstein, S.L. and Caves, C.M. (1994). Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72** 3439–3443.

[8] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton: Princeton Univ. Press. MR0016588

[9] Gibilisco, P. and Isola, T. (2003). Wigner–Yanase information in on quantum state space: The geometric approach. *J. Math. Phys.* **44** 3752–3762. MR2003930

[10] Gibilisco, P., Imparato, D. and Isola, T. (2009). Inequalities for quantum Fisher information. *Proc. Amer. Math. Soc.* **137** 317–327. MR2439456

[11] Guta, M. and Kahn, J. (2006). Local asymptotic normality for qubit states. *Phys. Rev. A* **73** 1–15. MR2229156

[12] Jencova, A. (2004). Geodesic distances on density matrices. *J. Math. Phys.* **45** 1787–1794. MR2054131

[13] Hayashi, M. (2002). Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation. *J. Phys. A* **35** 7689–7727. MR1947128

[14] Helstrom, C.W. (1967). Minimum mean-square error of estimates in quantum statistics. *Phys. Lett. A* **25** 101–102.

[15] Helstrom, C.W. (1976). *Quantum Detection and Estimation Theory*. New York: Academic Press.

[16] Holevo, A.S. (1982). *Probabilistic and Statistical Aspect of Quantum Theory*. Amsterdam: North-Holland. MR0681693

[17] Horn, R.A. and Johnson, C.R. (1991). *Topics in Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR1091716

[18] Luati, A. (2004). Maximum Fisher information in mixed state quantum systems. *Ann. Statist.* **32** 1770–1779. MR2089142

[19] Luati, A. (2008). A note on Fisher–Helstrom information inequality in pure state models. *Sankhyā* **70** 25–37. MR2507466

[20] Luo, S.L. (2003). Wigner–Yanase skew information vs. quantum Fisher information. *Proc. Amer. Math. Soc.* **132** 885–890. MR2019969

[21] Luo, S.L. (2006). Fisher information of wavefunctions: Classical and quantum. *Chinese Phys. Lett.* **23** 3127–3130.

[22] Nielsen, M.A. and Chuang, I.L. (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge Univ. Press. MR1796805

[23] Petz, D. (1996). Monotone metrics on matrix spaces. *Linear Algebra Appl.* **244** 81–96. MR1403277

[24] Petz, D. and Sudar, C. (1996). Geometry of quantum states. *J. Math. Phys.* **37** 2662–2673. MR1390228

[25] Petz, D. (2008). *Quantum Information Theory and Quantum Statistics*. Berlin: Springer. MR2363070

[26] Wigner, E.P. and Yanase, M.M. (1963). Information contents of distributions. *Proc. Natl. Acad. Sci. USA* **49** 910–918. MR0151127

[27] Yuen, H. and Lax, M. (1973). Multiple-parameter quantum estimation and measurement of nonselfadjoint observables. *IEEE Trans. Inform. Theory* **IT-19** 740–750. MR0366507

Received July 2009 and revised April 2010